Do Japanese Stock Prices Reflect Fundamentals?

Wenjuan Chen*
Free University of Berlin
Email:wenjuan.chen@fu-berlin.de

Anton Velinov
European University Institute
Email:anton.velionv@eui.eu

November 30, 2011

Abstract

This paper investigates to what extent the fundamentals of its real economy are reflected in the stock prices of Japan. A bivariate Structural Vector Autoregressive (SVAR) model with long run restrictions is utilized to identify the real activity shock and its contribution to the volatility of stock prices. Results show that the stock price in Japan has been priced above its fundamental component since the 1980s. Before the Japanese asset price bubble collapsed, less than a quarter of the variability in Japanese stock prices can be explained by the real activity shocks.

Keywords: Stock price, real activity, SVAR

JEL classifications: G12, E23

*Support by the Deutsche Forschungsgemeinschaft through the CRC 649 'Economic Risk' is greatly acknowledged.

1 Introduction

Many economists believe that the security prices fully reflect true fundamentals. However, the empirical evidence on the efficient market hypothesis is mixed. Some studies claim that the stock prices are closely related with real activities. Both James, Koreisha, and Partch (1985) and Chen, Roll, and Ross (1986) find that industrial production is one of the most significant factors in explaining share prices. Gjerde and Saettem (1999) analyze the relations among stock returns and macroeconomic variables for the Norwegian economy. They find that the real interest rate and the real activities are important variables for explaining returns.

Others argue that the the relationship between stock prices and fundamentals is rather weak. Binswanger (2004) and Groenewold (2004) both utilize a bivariate SVAR model estimated on data of the stock price and the industrial production. Their results provide evidence that stocks are priced substantially above their fundamentals from the mid 1990s. Cheung and Ng (1998) demonstrate similar findings with regression on earnings and dividends as fundamentals.

The purpose of this paper is to investigate the extent that Japanese stock price reflects real activities. A structural VAR framework is employed to study the relation between these two variables. Thus the correct identification of fundamental shocks is crucial. We also split the sample into pre the collapse of the Japanese asset price bubble and post the collapse, so as to find out whether the connection between stock prices and fundamentals has changed.

In order to disentangle the fundamental shocks and non-fundamental shocks, a long-run identification strategy is employed. Nevertheless, the appropriateness of the structural information used for identification could be questionable, as pointed by Uhlig (2005). Therefore we test whether the structural restrictions imposed on our Structural VAR model is appropriate or not applying a Markov regime switching framework following Lanne, Lütkepohl, and Maciejowska (2010).

Our main results demonstrate that nonfundamental shocks have been a main driving factor behind the fluctuations in the real stock price, even though the linkage between the Japanese stock price and its real activities has been strengthened since the Japanese asset price bubble collapsed.

This paper is structured as follows: Section 2 introduces the basic structural VAR model and the Markov switching VAR model used for validating identification. Section 3 describe our data and the validation of our identification strategy. Section 4 discusses the empirical application of the above model on the relation between the Japanese stock price and its real economy. Section 5 concludes.

2 The Model Setup

2.1 The SVAR Model and the Identification Strategy

Following Binswanger (2004) and Groenewold (2004), the following bivariate VAR model is chosen to study the interdependence of stock prices and the real activities:

$$\Delta x_t = \nu + A_1 \Delta x_{t-1} + A_2 \Delta x_{t-2} + \cdots + A_p \Delta x_{t-p} + u_t, \tag{1}$$

where ν is a $K \times 1$ vector of constants, K being the number of endogenous variables. Δx_t is a $K \times 1$ vector of the endogenous variables, in this case, $\Delta x_t = [\Delta y_t, \Delta s_t]'$ where the two variables $[y_t, s_t]'$ represent the log of industrial production and the log of real stock prices, and Δ is the first difference operator. A_i's are $K \times K$ parameter matrices, $i = 1, \ldots, p$ and u_t is a $K \times 1$ vector of unobservable error terms with $E[u_t] = 0$ and $E[u_t u_t'] = \Sigma_u$, not necessarily diagonal.

2.2 Identifying restrictions

Reduced form VAR alone can not disentangle the shocks hitting the system. One popular way to identify the shocks is to impose restrictions on the long-run impact matrix, Ψ as in Blanchard and Quah (1989). Following Binswanger (2004) and Groenewold (2004), we set the upper right element,

$\Psi_{1,2}$ of the long-run impact matrix to zero making it lower triangular.

$$\Psi = \begin{bmatrix} * & 0 \\ * & * \end{bmatrix} \tag{2}$$

where $*$ can take on any value.

Consequently the structural shocks, $\varepsilon_t = [\varepsilon_t^F, \varepsilon_t^{NF}]'$ can be interpreted as fundamental and non-fundamental shocks respectively. Hence, it is assumed that a fundamental shock can have a permanent effect on the real economy and on the stock market, while a non-fundamental shock can only have a transitory effect on the real economy and a permanent effect on the stock price.

2.3 Validating Identifying Restrictions

Various literature such as Uhlig (2005) have criticized that the structural restrictions could be too restrictive. Following Lanne, Lütkepohl, and Maciejowska (2010), we use a Markov Switching model to validate our identification strategy. This model allow for non-normality and heteroscedasticity of the residuals as follows:

$$\Delta x_t = \nu + A_1 \Delta x_{t-1} + A_2 \Delta x_{t-2} + \cdots + A_p \Delta x_{t-p} + u_t | s_t. \tag{3}$$

where the distribution of the residuals is assumed to be governed by a Markov process, s_t and it is assumed that the residuals are normally distributed in certain state, i.e., $u_t | s_t \sim N(0, \Sigma_{s_t})$

The discrete stochastic process s_t is assumed to take on M regimes with transition probabilities given by

$$p_{ij} = P(s_t = j | s_{t-1} = i), \quad i, j = 1, \dots, M$$

and an $M \times M$ matrix of transitional probabilities

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1M} \\ p_{21} & p_{22} & \cdots & p_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M1} & p_{M2} & \cdots & p_{MM} \end{bmatrix}. \tag{4}$$

Note that the probabilities add up to one row-wise, hence $p_{iM} = 1 - p_{i1} - p_{i2} - \cdots - p_{iM-1}, \quad i = 1, \ldots, M$.

Statistical information can be obtained from the above Markov switching VAR model Equation 3 to test the appropriateness of structural restrictions, given that the covariance matrices could be uniquely decomposed in the following way:

$$\Sigma_1 = BB', \quad \Sigma_2 = B\Lambda_2 B', \quad \ldots \quad \Sigma_M = B\Lambda_M B', \tag{5}$$

where B is the contemporaneous impact matrix and $\Lambda_i =$diag$(\lambda_{i1}, \lambda_{i2}, \ldots, \lambda_{iK}), \quad i = 2, \ldots, M$.

In order to validate the above proposed long run restrictions, it is necessary to compare the estimates from the unrestricted and the restricted Markov switching VAR models. Details of the Expectation-Maximization algorithm used for solving these models are given in the Appendix. With the log-likelihoods of both models, a simple likelihood ratio test will be used to determine their similarity. The next section describes the data and our empirical results on the stock market of Japan.

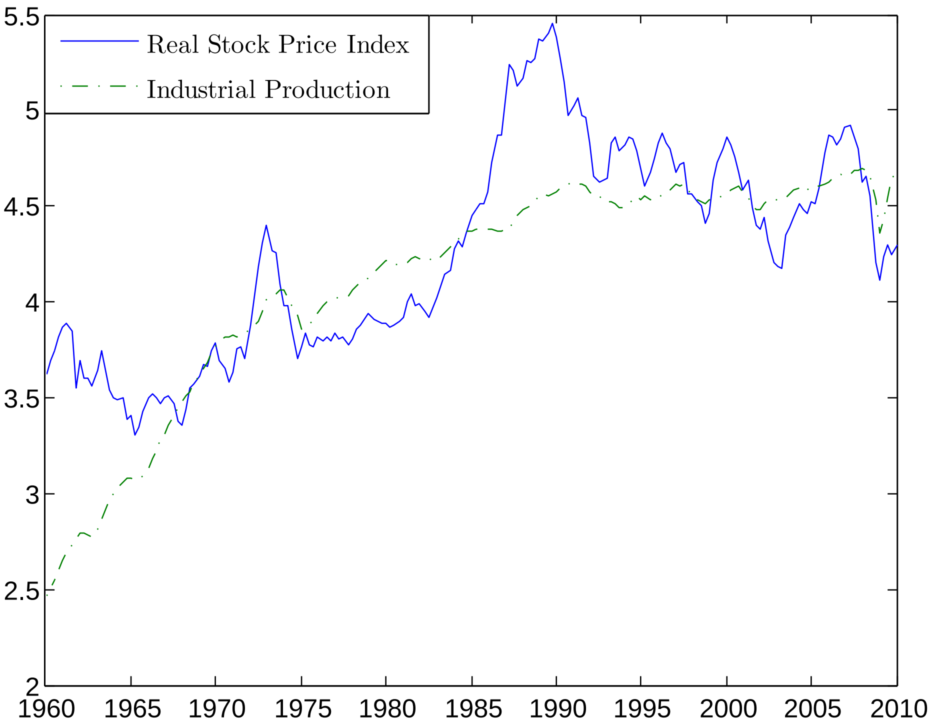
3 The Data and the Structural Identification

3.1 The Data

Most of our data are obtained from the International Financial Statistics (IFS) of the IMF. The series consist of industrial production, a stock price index and Consumer Price Index (CPI), all of which are normalized to a base year of 2005. The stock price series is converted to real terms by dividing by the Consumer Price Index, hence the CPI series is not used directly in the analysis. Industrial production is seasonally adjusted. In addition all series are in logs. The data range is quarterly from 1960:I-2010:I. Note that the financial crisis is included.

ADF unit root tests show that the levels series for all countries are of order $I(1)$ meaning that the first differences are stationary, as can be seen from Figure 1. When testing for cointegrating relationships for the unrestricted levels VAR model using a constant and trend term, both the Saikkonen and Lütkepohl (S&L) and the Johansen test clearly cannot reject the null hypothesis of no cointegrating relations at usually the 1% level. There seems to be no cointegrating relations among the variables. Hence it is justified to use the standard VAR model in first differences depicted in equation (1).

Figure 1: Industrial production and real stock price series in log levels



The optimal lag length is chosen according to the Akaike Information Criterion (AIC) when using a maximum lag length of 8 and an intercept term. Although the AIC tends to sometimes overestimate the lag length, other criteria such as the Hannan-Quinn (HQ) and the Schwarz Criterion (SC) tend to underestimate it and sometimes favor a model without any lags which is not desirable. Taking this into consideration we decide to use a one-lag model for Japan.

Table 1: Parameter estimates and standard errors for Japan Unrestricted MS-VAR

	estimates	standard errors
λ_{21}	2.912	1.186
λ_{22}	2.411	0.802
λ_{31}	55.637	41.765
λ_{32}	3.109	2.046
p_{21}	0.963	0.031
p_{22}	0.812	0.092
p_{33}	0.866	0.479
p_{12}	0.037	0.027
p_{21}	0.155	0.106
p_{32}	0.134	0.347

3.2 Are Identifying Restrictions Appropriate?

Based on the model selection criteria developed in Psaradakis and Spagnolo (2003) and Psaradakis and Spagnolo (2006), we select a three-state one-lag Markov switching structural VAR model for Japan. The smoothed probabilities for this model are depicted in Figures 4. The volatility is increasing in states. The third state is the most volatile one, which happens around the the time of the late 2000s financial crisis. The period of relative clam known as the great moderation is also visible.

The parameter estimates for Japan are found in Tables1. Is our decomposition of covariance matrix unique? It seems that the parameters of the Λ_i, $i = 2, \ldots, M$ matrices are quite different, however their standard errors are also quite large in some cases. In order to test whether they are indeed different, Wald and Likelihood Ratio (LR) tests are performed, the results of which are shown in Tables 2 with p-values obtained from a χ^2 distribution. From the Wald test results for the unrestricted models reported in Table 2, the null hypotheses of identical parameters of the Λ_2 matrix can be rejected at the 5% level. This is even more strongly reaffirmed when looking at the LR test results. Hence, the decomposition in (5) is unique up to sign

Table 2: Is the Decomposition Unique?

Model	Wald tests				LR tests	
	Unrestricted		Restricted		Unrestricted	
	test value	p-value	test value	p-value	test value	p-value
	1.610	0.447	2.147	0.342	8.281	0.016

Table 3: Are the Identifying Restrictions Appropriate?

	H_0: Restricted	H_1: Unrestricted
Country	test value	p-value
Japan	2.449	0.118

changes in the B matrix.

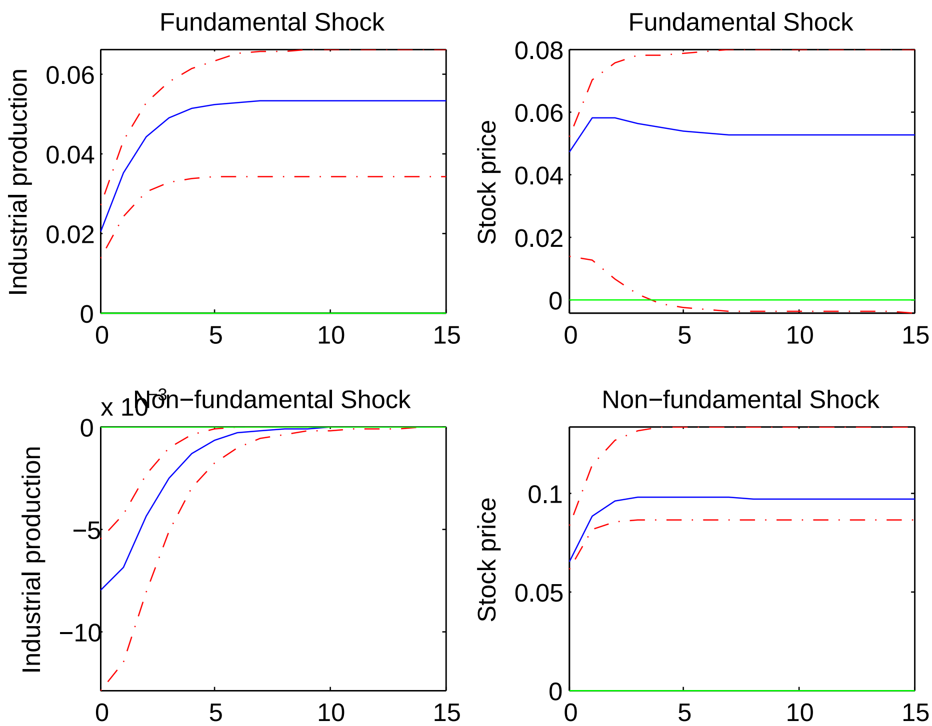
Based on the above unique decomposition, we now turn to a likelihood ratio test to find out whether the long run restriction imposed on our SVAR model is supported by the data. Table 3 displays the LR test results on the identification restriction used on the long-run impact matrix Ψ in (2). The restricted model is supported at the 5% level for both Japan and the US. Hence, the characterization of the shocks as fundamental and non-fundamental in our model is valid.

4 Stock Price Pre- and Post- the Collapse of Japanese Asset Price Bubbles

In the above section we have validated the long run restriction imposed to disentangle fundamental shocks and nonfundamental shocks. Based on the appropriate structural identification, now we conduct a standard SVAR analysis to determine whether stock prices are driven by the fundamentals in Japan. Figure 2 shows the accumulated impulse responses for Japan.

Confidence intervals are according to the fixed design wild bootstrap as in Goncalves and Kilian (2004) for the 95% level. The first row in the figure displays the response of industrial production and the stock price to a fundamental shock.

Figure 2: Accumulated Impulse Responses. Confidence intervals are according to fixed design wild bootstrap at the 95% level.



Next, Forecast Error Variance Decompositions (FEVD's) of the real stock price are shown in Table 4. The fundamental shock appears to explain relatively little of the forecast error variance of the real stock price for up to 20 quarters ahead. After 10 quarters already the result stabilizes and the proportion of the forecast error explained by each shock stays the same. The results of the FEVD's would point to a weak relationship between real stock prices and industrial production, this lends evidence to support a bubble component in stock prices.

What would the Japanese stock price look like if it has only been affected by fundamental shocks? Following the method proposed by Burbidge and

Table 4: FEVD's of the real stock price

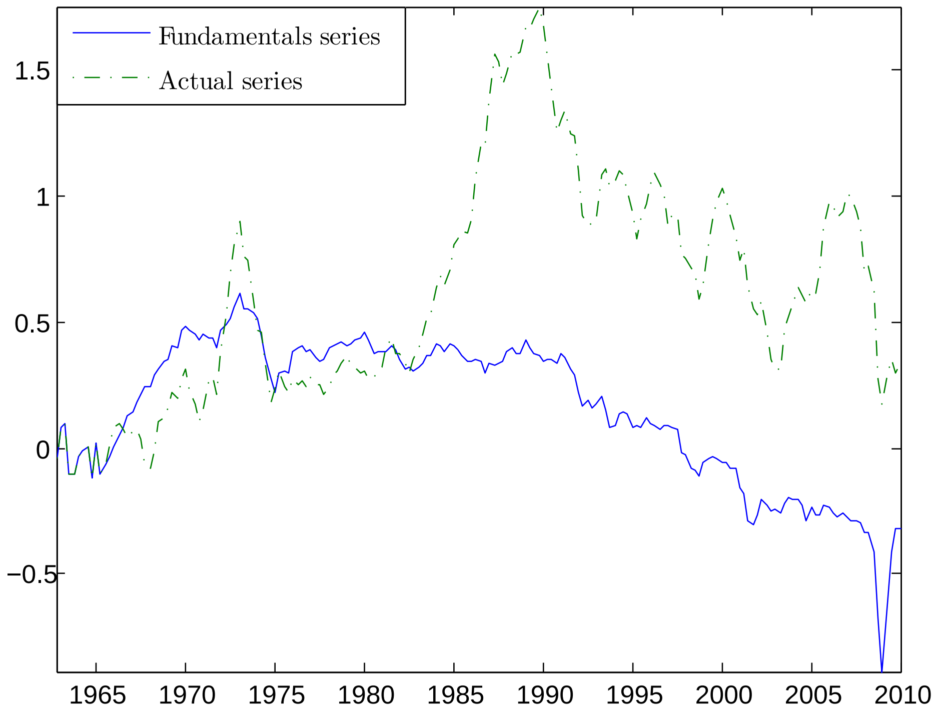
	Percentage of variance attributable to:			Percentage of variance attributable to:	
	Fundamental shock	Non-fundamental shock		Fundamental shock	Non-fundamental shock
	1960-1989			1990-2010	
1	23	77	1	46	54
2	22	78	2	44	56
3	22	78	3	43	57
4	23	77	4	43	57
5	23	77	5	43	57
10	23	77	10	43	57
15	23	77	15	43	57
20	23	77	20	43	57

Harrison (1985), we set the value of nonfundamental shocks as zero and simulate the historical values of the Japanese stock price in the presence of only real activity shocks from 1965 to 2010.

Graphs of historical decomposition for the Japanese stock price are displayed in Figure 3. From the early 1970s until the mid 1980s the fundamentals series is moving quite closely with the actual series. However, the real stock price started floating substantially above the fundamental afterwards, reaching its peak shortly before 1990. The Japanese stock price has since stayed far above its fundamentals for the whole period known as the lost decades. This confirms the findings by Chung and Lee (1998) and Lee (1995).

To sum up, the above analysis suggests a weak relation between the stock prices in Japan and its real economy from 1960 to 2010. Especially, the graph of Historical Decomposition provide evidence of bubbles in share prices from the mid 1980s to the beginning of 1990s. After the collapse of the Japanese asset price bubble, the linkage between the stock price and fundamentals became stronger, however, the late 2000s financial crisis posted a new challenge.

Figure 3: Historical Decompositions (HD's) for Japanese Stock Price



Notes: This graph depicts historical decomposition for the Japanese stock price from 1965 till 2010. The fundamentals series refers to the decomposed series in which only fundamental shocks can influence stock prices, i.e. with the non-fundamental shocks set to zero.

5 Conclusion

This paper has investigated how much real activity shocks can explain the fluctuations in the Japanese stock price in a bivariate SVAR model. We found that the Japanese stock price has been priced substantially above fundamentals even after the collapse of the Japanese asset price bubbles, though the proportion of nonfundamental-driven variance after lessened after the collapse.

Our results shed doubts on the efficient market hypothesis. In recent years, work by psychologists have identified overconfidence and social contagion as causes of stock market bubbles. Interdisciplinary studies with anthropologists and psychologists could help to explain why Japanese stock price has been overvalued for more than two decades.

References

BINSWANGER, M. (2004): "How important are fundamentals? Evidence from a structural VAR model for the stock markets in the US, Japan and Europe," *Journal of International Financial Markets, Institutions & Money*, 14(2), 185–201.

BLANCHARD, O., AND D. QUAH (1989): "The dynamic effects of aggregate demand and supply disturbances," *The American Economic Review*, 79(4), 655–673.

BURBIDGE, J., AND A. HARRISON (1985): "An historical decomposition of the great depression to determine the role of money," *Journal of Monetary Economics*, 16(1), 45–54.

CHEN, N., R. ROLL, AND S. ROSS (1986): "Economic forces and the stock market," *Journal of Business*, 59(3), 383–403.

CHEUNG, Y., AND L. NG (1998): "International evidence on the stock market and aggregate economic activity," *Journal of Empirical Finance*, 5(3), 281–296.

CHUNG, H., AND B. LEE (1998): "Fundamental and nonfundamental components in stock prices of Pacific-Rim countries," *Pacific-Basin Finance Journal*, 6(3-4), 321–346.

GJERDE, Ø., AND F. SAETTEM (1999): "Causal relations among stock returns and macroeconomic variables in a small, open economy," *Journal of International Financial Markets, Institutions and Money*, 9(1), 61–74.

GONCALVES, S., AND L. KILIAN (2004): "Bootstrapping autoregressions with conditional heteroskedasticity of unknown form," *Journal of Econometrics*, 123(1), 89–120.

Groenewold, N. (2004): "Fundamental share prices and aggregate real output," *Applied Financial Economics*, 14(9), 651–661.

Hamilton, J. (1994): *Time series analysis*. Princeton Univ Pr.

James, C., S. Koreisha, and M. Partch (1985): "A VARMA analysis of the causal relations among stock returns, real output, and nominal interest rates," *Journal of Finance*, 40(5), 1375–1384.

Krolzig, H. (1997): "Markov-Switching Vector Autoregressions–Modelling," *Statistical Inference and Application to Business Cycle Analysis*.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010): "Structural Vector Autoregressions with Markov Switching," *Journal of Economic Dynamics and Control*, 34(2), 121–131.

Lee, B. (1995): "The response of stock prices to permanent and temporary shocks to dividends," *Journal of Financial and Quantitative Analysis*, pp. 1–22.

Psaradakis, Z., and N. Spagnolo (2003): "ON THE DETERMINATION OF THE NUMBER OF REGIMES IN MARKOV-SWITCHING AUTOREGRESSIVE MODELS," *Journal of time series analysis*, 24(2), 237–252.

——— (2006): "Joint determination of the state dimension and autoregressive order for models with Markov regime switching," *Journal of Time Series Analysis*, 27(5), 753–766.

Uhlig, H. (2005): "What are the effects of monetary policy on output? Results from an agnostic identification procedure," *Journal of Monetary Economics*, 52(2), 381–419.

Appendix

This is a technical appendix explaining the EM algorithm used in this paper in more detail. The following is based on Krolzig (1997).

Starting with the regression equation

$$\Delta x = (\bar{Z} \otimes I_K)\beta + u,$$

where Δx is a $(TK \times 1)$ vector or the vectorization of $\Delta X = [\Delta x_1, \ldots, \Delta x_T]$, and where T is the sample size and K the number of variables. Here $\bar{Z} = [\mathbf{1}_T, \Delta X_{-1}, \ldots, \Delta X_{-p}]$, where $\mathbf{1}_T$ is a $(T \times 1)$ vector of ones and $\Delta X_{-i} = [\Delta x_{1-i}, \ldots, \Delta x_{T-i}]'$ is a $(T \times K)$ matrix of lagged regressors, for $i = 1, \ldots, p$ and p being the number of lags of the MS-VAR model. The $(K(Kp + 1) \times 1)$ vector β contains the vectorized intercept and slope parameters, i.e. $\text{vec}[\nu, A_1, \ldots, A_p]$ as defined in (1). Finally u is the $(TK \times 1)$ vectorization of the matrix of residuals, $U = [u_1, \ldots, u_T]'$, where the distribution of each residual, $u_i, i = 1, \ldots, T$ is given according to (3).

The EM algorithm is initiated by defining the starting values of the intercept, slope and contemporaneous impact matrix, B parameters as well as the transition probabilities and initial states. For the intercept and slope parameters the starting values are given by $\beta_0 = [\bar{Z}'\bar{Z} \otimes I_K]^{-1}(\bar{Z}' \otimes I_K)\Delta x$. The initial value of the contemporaneous impact matrix is $B_0 = (UU'/T)^{1/2}$, where U is obtained from $u = \Delta x - (\bar{Z} \otimes I_K)\beta_0$. The transition probabilities are set at $P_0 = \mathbf{1}_M \mathbf{1}_M'/M$, where $\mathbf{1}_M$ is an $(M \times 1)$ vector of ones and M are the number of states in the model. The initial states (defined below) are defined as $\xi_{0|0} = \mathbf{1}_M/M$. Finally, the starting values of the covariance matrices need to be determined as defined in the decomposition in (5). This is done by setting the values of the Λ_i matrices, $i = 2, \ldots, M$. I use a loop of different starting values for these matrices by starting with $\Lambda_2 = 2 * I_K, \ldots, \Lambda_M = 2^{M-1} * I_K$ and replacing the 2 with higher values and in the end seeing which starting value gives the highest log-likelihood.

The vector of conditional probabilities for the unobserved states is denoted as $\hat{\xi}_{t|t}$ and it indicates the probability of a given state in a given time period conditional on all observations up to time period t, ΔX_t and all intercept, slope, covariance parameters and transition probabilities stored in, θ. Hence

$$\hat{\xi}_{t|t} = \begin{bmatrix} P(s_t = 1|\Delta X_t, \theta) \\ P(s_t = 2|\Delta X_t, \theta) \\ \vdots \\ P(s_t = M|\Delta X_t, \theta) \end{bmatrix}. \tag{6}$$

It is also necessary to define the conditional densities of an observation given a particular state, all past observations and θ as

$$\eta_t = \begin{bmatrix} P\Delta x_t|s_t = 1, \Delta X_{t-1}, \theta) \\ P\Delta x_t|s_t = 2, \Delta X_{t-1}, \theta) \\ \vdots \\ P(\Delta x_t|s_t = M, \Delta X_{t-1}, \theta) \end{bmatrix} = \begin{bmatrix} \frac{1}{2\pi|\Sigma_1|^{1/2}}\exp\Big\{-\frac{u_t'\Sigma_1^{-1}u_t}{2}\Big\} \\ \frac{1}{2\pi|\Sigma_2|^{1/2}}\exp\Big\{-\frac{u_t'\Sigma_2^{-1}u_t}{2}\Big\} \\ \vdots \\ \frac{1}{2\pi|\Sigma_M|^{1/2}}\exp\Big\{-\frac{u_t'\Sigma_M^{-1}u_t}{2}\Big\} \end{bmatrix}. \tag{7}$$

5.1 Expectation Step

Now follows the expectation step where the filtered probabilities from (6) are calculated as

$$\hat{\xi}_{t|t} = \frac{\eta_t \odot \hat{\xi}_{t|t-1}}{\mathbf{1'}(\eta_t \odot \hat{\xi}_{t|t-1})}, \tag{8}$$

and

$$\hat{\xi}_{t|t-1} = P'\hat{\xi}_{t-1|t-1}, \tag{9}$$

for $t = 1, \ldots, T$. This generates an $(M \times 1)$ vector of conditional probabilities for each time period. Here $\odot$ denotes element-by-element multiplication and P is defined as in (4). Next using the values of the filtered probabilities, the smoothed probabilities, $P(s_t = i|\Delta X_T, \theta), i = 1, \ldots, M$ are estimated as

$$\hat{\xi}_{t|T} = [P(\hat{\xi}_{t+1|T} \oslash \hat{\xi}_{t+1|t})] \odot \hat{\xi}_{t|t}, \tag{10}$$

for $t = T-1, \ldots, 0$. The symbol $\oslash$ denotes element-by-element division. Note that the filtered probabilities from the current iteration are used to estimate the smoothed probabilities.

5.2 Maximization Step

After the expectation step in the maximization step first the vector of transition probabilities $\hat{\rho}$ is estimated as

$$\hat{\rho} = \hat{\xi}^{(2)} \oslash (\mathbf{1}_M \otimes \hat{\xi}^{(1)}), \tag{11}$$

where $\hat{\xi}^{(2)} = \sum_{t=0}^{T-1} \hat{\xi}_{t|T}^{(2)}$ and

$$\hat{\xi}_{t|T}^{(2)} = \text{vec}(P) \odot \left[\left(\hat{\xi}_{t+1|T}^{(1)} \oslash \hat{\xi}_{t+1|t}^{(1)} \right) \otimes \hat{\xi}_{t|t}^{(1)} \right],$$

for $t = 0, \dots, T-1$. Here $\otimes$ denotes the Kronecker product. Finally, $\hat{\xi}_{t|T}^{(1)}$ is the vector of smoothed probabilities from (8) and $\hat{\xi}_{t|t}^{(1)}$ is the vector of filtered probabilities from (6). Also note that $\hat{\xi}^{(1)} = (\mathbf{1}'_M \otimes I_M)\hat{\xi}^{(2)}$, where $\mathbf{1}_M$ is an $(M \times 1)$ vector of ones and I_M is the $(M \times M)$ identity matrix.

The B and Λ matrices are then estimated by optimizing

$$\begin{aligned} l(B, \Lambda_2, \dots, \Lambda_M) &= T \log|\det(B)| + \frac{1}{2}\text{tr}\left((BB')^{-1}\hat{U}\hat{\Xi}_1\hat{U}' \right) \\ &+ \sum_{m=2}^{M} \left[\frac{\hat{T}_m}{2}\log(\det(\Lambda_m)) + \frac{1}{2}\text{tr}\left((B\Lambda_m B')^{-1}\hat{U}\hat{\Xi}_m\hat{U}' \right) \right] \end{aligned} \tag{12}$$

where $\hat{U}$ is obtained from $\hat{u} = \Delta x - (\bar{Z} \otimes I_K)\hat{\beta}$, $\hat{\Xi}_m = \text{diag}(\hat{\xi}_{m1|T}, \dots, \hat{\xi}_{mT|T})$, the smoothed probabilities of regime m and $\hat{T}_m = \sum_{t=1}^{T} \hat{\xi}_{mt|T}$ is a summation of the smoothed probabilities. To avoid singularity a lower bound of 0.001 is imposed on the diagonal elements of the $\Lambda_m, m = 2, \dots, M$ matrices. The updated covariance matrices are given from the decomposition

$$\hat{\Sigma}_1 = \hat{B}\hat{B}', \quad \hat{\Sigma}_2 = \hat{B}\hat{\Lambda}_2\hat{B}', \quad \dots \quad \hat{\Sigma}_M = \hat{B}\hat{\Lambda}_M\hat{B}'.$$

Next the intercept and slope parameters are obtained as

$$\hat{\beta} = \left[\sum_{m=1}^{M} (\bar{Z}'\hat{\Xi}_m\bar{Z}) \otimes \hat{\Sigma}_m^{-1} \right]^{-1} \left[\sum_{m=1}^{M} (\bar{Z}'\hat{\Xi}_m) \otimes \hat{\Sigma}_m^{-1} \right] \Delta x. \tag{13}$$

Note, that to estimate $\hat{\beta}$ the covariances of the previous iteration were used. These parameters are then plugged back into (12) and new estimates of the covariance matrices are obtained which are then used in (13). All this is iterated until convergence. The convergence criteria used is the absolute change in the log-likelihood given in (12), i.e.

$$\Delta = |l(\theta^{j+1}|\Delta X_T) - l(\theta^j|\Delta X_T)|, \tag{14}$$

where $l(\bullet)$ is the log-likelihood and θ^j denotes the parameters of the j-th iteration. Convergence is satisfied when $\Delta \leq 10^{-6}$ or after a specified maximum number of iterations.

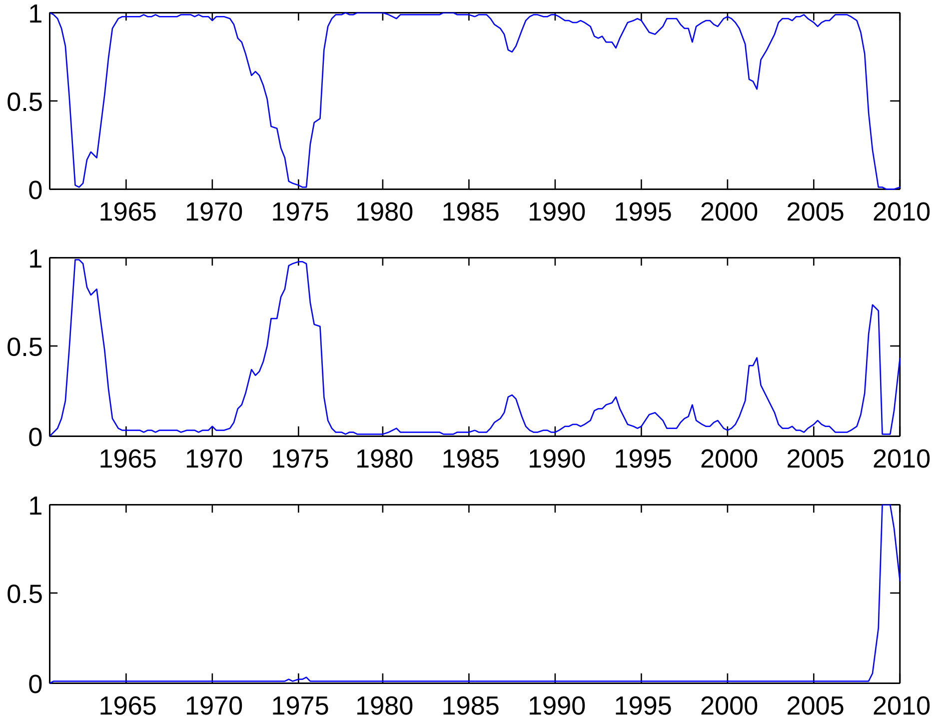
The EM algorithm terminates as well after a similar convergence criteria as in (14). As shown in Hamilton (1994) the log-likelihood is given by $\log(\mathbf{1'}(\eta_t \odot \hat{\xi}_{t|t-1}))$.

The restricted MS-SVAR model is estimated in a similar way, recall that the long-run impact matrix, Ψ is related to the B matrix by $\Psi = A(1)^{-1}B$.

5.3 Standard Errors

Once the EM algorithm has converged and the point estimates of the parameters are obtained it is necessary to calculate their standard errors in order to carry out statistical tests. The optimal values of $P, \beta, B, \Lambda_m, m = 2, \ldots, M$ and $\xi_{0|0}$ are used in $\log(\mathbf{1'}(\eta_t \odot \hat{\xi}_{t|t-1}))$. Standard errors are then obtained by the inverse of the negative of the Hessian matrix.

Figure 4: Smoothed probabilities for Japan



Notes: This graph depicts the smoothed probabilities estimated from the Markov Switching VAR model with three states and one lag. The top graph shows the probability of the system being in a low-volatility regime. The graph in the middle represents the probability of being in a medium-volatility regime, while the bottom graph represents the probability of being in a high-volatility regime.